Plasma-glycated CD59 as an early biomarker for gestational diabetes mellitus: prospective cohort study protocol

Chloe Andrews, Michelle Toth-Castillo, Huseyin Aktas, Miguel-Angel Luque Fernandez, Steven Koon Wong, Sarbattama Sen, Jose Halperin

ABSTRACT:

Introduction The significant maternal and neonatal outcomes of gestational diabetes mellitus (GDM) make it a major public health concern. Mothers with GDM are at greater risk of pregnancy complications and their offspring are at higher risk of diabetes and obesity. Currently, GDM is diagnosed with glucose load methods which are time-consuming and inconvenient to administer more than once during pregnancy; for this reason, there is a recognised need for a more accurate and simpler test for GDM. Previous studies indicate that plasma-glycated CD59 (pGCD59) is a novel biomarker for GDM. We present here the protocol of a prospective cohort study designed to (1) determine the accuracy of pGCD59 as an early, first trimester predictor of GDM and gestational impaired glucose tolerance and (2) assess the associations between pGCD59 levels and adverse maternal and neonatal outcomes.

Methods and analysis We will obtain discarded plasma samples from pregnant women at two time points: first prenatal visit (usually <14 weeks gestation) and gestational weeks 24–28. A study-specific medical record abstraction tool will be used to obtain relevant maternal and neonatal clinical data from the EPIC clinical database. The prevalence of GDM will be determined using standard care glucose load test results. We will determine the sensitivity and specificity of pGCD59 to predict the diagnosis of GDM and gestational impaired glucose tolerance, as well as the associations between levels of pGCD59 and the prevalence of maternal and neonatal outcomes.

Ethics and dissemination This study has been approved by the Mass General Brigham Institutional Review Board (protocol 2011P002254). The results of this study will be presented at international meetings and disseminated in peer-reviewed journals.

INTRODUCTION

Gestational diabetes mellitus (GDM), defined as glucose intolerance that first occurs or is identified in pregnancy, is estimated to affect 7%–10% of pregnancies in the USA every year. This condition represents a major risk factor for both maternal and infant adverse pregnancy outcomes: for mothers with GDM, there is a greater risk of pre-eclampsia, caesarean section and preterm delivery; for their infants, there is a greater risk of macrosomia, large-for-gestational age (LGA), neonatal hypoglycaemia and future development of obesity, diabetes and cardiovascular disease. Treatment of GDM, as well as weight control following diagnosis, reduces the incidence of some adverse outcomes, strongly underscoring the importance of accurate and minimally burdensome universal screening, as well as early diagnosis.

Current guidelines recommend GDM screening at 24–28 weeks gestation. However, markers of glycaemic load used in the non-pregnant population, such as fasting plasma glucose, haemoglobin A1c (HbA1c) and fructosamine, notably fail as screening or diagnostic tests for GDM. HbA1c, the standard of care test for management of individuals with diabetes, provides an integrated measure of glycaemic load over a period of 6–8 weeks.
but has poor sensitivity for the screening/diagnosis of GDM due in part to the increased turnover rate of erythrocytes during pregnancy. Furthermore, the worsening insulin resistance in pregnancy is initially compensated for by an increase in insulin production, resulting in normal or close to normal fasting glucose levels.11 12 This also explains the low sensitivity of fasting plasma glucose, which provides a transient snapshot of glucose levels.12 Regarding A1c and fructosamine, four studies that evaluated several HbA1c thresholds using different diagnostic criteria for GDM failed to identify a clear pattern between A1c levels and probability of GDM diagnosed by Carpenter-Coustan (C&C) criteria.13 Similarly, several human studies demonstrated that fructosamine, another marker that integrates glycaemic load over time, has very limited value as a screening test for GDM.14 15

The low sensitivity of traditional glycaemic markers in pregnancy explains why glucose load tests (GLTs) are still universally used as the standard of care. A ‘two-step’ approach, consisting of a 1-hour 50-g GLT followed by a 3-hour 100-g oral glucose tolerance test (OGTT) for those who fail the GLT, is mostly used in the USA, whereas a ‘one-step’ approach, consisting of a 2-hour 75-g load, is used in many countries around the world.16 Both approaches are expensive, time consuming, unpleasant and have poor reproducibility on repeat testing of the same individual; moreover, their predictive value is low and differs according to ethnic origin.17 The failure rate of the OGTT to make a diagnosis of GDM is nearly 10%, with vomiting during the test being the major reason for this failure.18 In addition, the seminal Hyperglycemia and Adverse Pregnancy Outcome study has clearly documented a continuous positive association between increased values of maternal glycaemia and the risk of adverse maternal and fetal outcomes.4 As a consequence, there is currently debate as to the rationale of imposing an arbitrary dichotomous definition of ‘normal’ and ‘abnormal’ on gestational glucose tolerance. Indeed, empirical evidence accumulating over the last 20 years supports recognising an intermediate category of abnormal gestational glucose tolerance in women that falls in between the currently accepted ‘normal’ and ‘abnormal’ glucose thresholds. This intermediate category of glucose intolerance is also associated with considerable increases in risk of fetal macrosomia (27.5% vs 9.9% for intermediate vs normal, respectively), pre-eclampsia, caesarean-section delivery (40.0% vs 19.9% for intermediate vs normal, respectively), preterm delivery, neonatal jaundice and extended maternal and infant hospital length of stay.16–23

Together, the discussion above explains the need for a simpler, easier-to-use, cost-effective, sensitive and specific biomarker for detection of GDM.24 Prior evidence supports a strong link between the complement system, activity of the complement regulatory protein CD59 and the pathogenesis of vascular complications of diabetes.25–26 Complement activation ultimately leads to formation of the pore-forming membrane attack complex (MAC), the main effector of complement-mediated tissue damage, CD59, a key complement regulator that inhibits MAC formation,27 is universally expressed in mammalian cells. Glycation, the non-enzymatic attachment of glucose to protein amino groups, forms glycated CD59 (GCD59) and abrogates its activity as an inhibitor of MAC formation.28 29 Phospholipases shed both CD59 and GCD59 off cell membranes and release them into the circulation in soluble forms that can be measured in blood and other bodily fluids.

We have developed an assay to measure blood levels of GCD59 and demonstrated that these levels accurately distinguish individuals with and without diabetes (HbA1c >6.5% or <6%, respectively).30 Furthermore, we previously reported that plasma GCD59 (pGCD59) levels at 24–28 weeks gestation identified GDM with high specificity and sensitivity.31 Also, higher maternal pGCD59 levels were associated with a higher prevalence of LGA even after adjustment for confounders such as body mass index (BMI), maternal and gestational age, race/ethnicity and history of diabetes.31 The prevalence of LGA newborns was 4.3% in the lowest quartile and 13.5% in the highest quartile of pGCD59. Recognising the importance of early detection of GDM, we also assessed the utility of pGCD59 as an early predictor (<20 weeks gestation) of GDM in 693 women with obesity from the Vitamin D and Lifestyle Intervention study.32 33 We found that pGCD59 successfully identified early GDM with an area under the receiver operating characteristic curve of 0.86. In this study, a one-unit increase in maternal pGCD59 concentration was associated with 36% increased odds of delivering an LGA infant.33

The assay we have developed uses two specific monoclonal antibodies in a sandwich ELISA; it is highly reproducible and measures serum and/or pGCD59 in a highly specific and sensitive manner. Indeed, assuming a basal level of CD59 glycation of ~5% (based on the basal glycation level of HbA1c in normoglycaemic individuals), and that the concentration of CD59 in human serum is ~100–150 ng/mL,34 the detection of GCD59 in individuals without diabetes implies that the assay can detect GCD59 in the low picomolar range.

We report here the protocol of a human study designed to assess the accuracy of pGCD59 as an early, first-trimester (<14 weeks gestation) predictor of GDM. Secondary objectives of this study are (1) to evaluate the potential association of maternal pGCD59 levels with adverse maternal and neonatal outcomes usually associated with GDM and (2) to assess the association of pGCD59 with the prevalence of gestational impaired glucose tolerance (GIGT), a form of glucose intolerance that does not meet American College of Obstetricians and Gynecologists criteria for GDM.

**METHODS AND ANALYSIS**

**Human subjects, study design and sample acquisition**

This is a prospective cohort study of pregnant women and their infants who receive prenatal care and deliver
at Brigham and Women’s Hospital (BWH) in Boston, Massachusetts. Patients who have consented to make their excess samples and associated clinical data available to researchers are included in the Partner’s Crimson Biobank. We will obtain excess samples from the biobank based on study-specific criteria. The clinical data will be extracted from the electronic medical records (EMRs) of patients who have consented to make their deidentified samples available to researchers.

**Sample selection**

Eligible plasma samples from donors meeting the criteria displayed in box 1 will be separated by the Crimson Biobank from tubes routinely used to measure HbA1c at the first prenatal visit.

**Sample size calculation**

We estimate that including 500 women diagnosed with GDM at 24–28 weeks gestation, diagnosed based on current standard of care (C&C criteria), will provide

---

**Box 1 Eligibility criteria for plasma-glycated CD59 study**

- Consented to participate in Crimson Biobank.
- Discarded clinical plasma sample at <14 weeks gestation.
- ≥18 and ≤40 years old.

---

**Table 1 Maternal and infant data obtained or derived from EPIC and the Clarity report**

<table>
<thead>
<tr>
<th>Data category/time point</th>
<th>Data points collected</th>
<th>Data points derived</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics</td>
<td>Race, ethnicity, insurance, OB practice.</td>
<td>–</td>
</tr>
<tr>
<td>Obstetrics and medical history</td>
<td>Gravidity, parity, medical history (ICD10 codes, active/resolved), infertility diagnosis (ICD10 codes).</td>
<td>–</td>
</tr>
<tr>
<td>Prenatal visits (three time points: first (V1), second (V2), third (V3) trimester)</td>
<td>Date of visit, age (V1 only), height (V1 only), weight, GA, blood pressure, medications.</td>
<td>BMI</td>
</tr>
<tr>
<td>Prenatal laboratory studies (three time points: first, second, third trimester)</td>
<td>Date of laboratory studies, GA at laboratory studies, WBCs, haematocrit, platelets, neutrophils, lymphocytes, haemoglobin A1c.</td>
<td>–</td>
</tr>
<tr>
<td>Prenatal ultrasounds (three time points: first, second, third trimester)</td>
<td>Date of ultrasound, fetal measurements (EFW, femur length, abdominal circumference, biparietal diameter, head circumference, weight percentile).</td>
<td>IUGR</td>
</tr>
<tr>
<td>Glucose load test</td>
<td>Date of GLT, GLT result.</td>
<td>GDM (C&amp;C criteria)</td>
</tr>
<tr>
<td>Glucose tolerance test</td>
<td>Date of GTT, GTT results (fasting and 1, 2 and 3 hours).</td>
<td>GDM, GIGT (C&amp;C criteria)</td>
</tr>
<tr>
<td>Additional prenatal laboratory studies</td>
<td>Results and dates of 24-hour protein, random urine protein, random protein, random creatinine, TSH.</td>
<td>–</td>
</tr>
<tr>
<td>Pregnancy complications</td>
<td>Clinical diagnosis (ICD10 codes) of GDM, pregnancy-induced hypertension, pre-eclampsia, pre-eclampsia with severe features, HELLP syndrome, UTIs, chorioamnionitis, URIs, STDs, neural tube defects, congenital heart defects, abdominal wall defects.</td>
<td>–</td>
</tr>
<tr>
<td>Delivery</td>
<td>Date and time of delivery, maternal weight, MOD, GA, infant MRN, sex, GBS, antibiotics, Apgar scores (1 and 5 min), infant measurements at birth (birth weight, length and head circumference), infant weight at discharge.</td>
<td>Multiple gestation, gestational weight gain (V1 to delivery), ponderal index, macrosomia, LGA, SGA</td>
</tr>
<tr>
<td>Infant feeding</td>
<td>Maternal feeding intention, date and time of first feed, breast milk at first feed, feeding type at discharge.</td>
<td>–</td>
</tr>
<tr>
<td>Infant clinical data</td>
<td>Diagnoses (ICD10 codes) of LGA, SGA, fetal malformations, shoulder dystocia/birth injury, hyperbilirubinaemia, NH and NICU admission; NICU length of stay, BG values, dextrose containing IVs, O2 devices, bilirubin, phototherapy, haematocrit.</td>
<td>Lowest BG, BG &lt;45 mg/dL</td>
</tr>
<tr>
<td>Postpartum maternal visit</td>
<td>Date of visit, weight, blood pressure, GTT results, infant feeding method.</td>
<td>Postpartum weight loss</td>
</tr>
</tbody>
</table>

BG, blood glucose; BMI, body mass index; C&C, Carpenter & Coustan; EFW, estimated fetal weight; GA, gestational age; GBS, group B streptococcus; GDM, gestational diabetes mellitus; GIGT, gestational impaired glucose tolerance; GLT, glucose load test; GTT, glucose tolerance test; ICD, International Classification of Diseases; IUGR, intrauterine growth restriction; LGA, large for gestational age; MOD, mode of delivery; MRN, medical record number; NH, neonatal hypoglycaemia; NICU, neonatal intensive care unit; OB, obstetrics; SGA, small for gestational age; STDs, sexually transmitted diseases; TSH, thyroid stimulating hormone; UTIs, urinary tract infections; WBCs, white blood cells.
sufficient power to assess the accuracy of first trimester pGCD59 to identify women with GDM. In 2016, the BWH clinical laboratory received approximately 6000 plasma samples from women undergoing a GLT in the course of their prenatal care. Of these women, 1156 (19.6%) had abnormal GLT results and 380 (6.4%) were subsequently diagnosed with GDM using C&C criteria. Based on these data, we conservatively estimate that to include 500 women with GDM we will need to measure first trimester pGCD59 in ~8000 women, given a prevalence of 5% of GDM in the BWH population.

Sample and data collection
After enrollment and collection of the first trimester plasma sample, a follow-up plasma sample will be collected for each subject at pregnancy weeks 24–28, when the GLT is performed. Both of these plasma samples (approximately 0.5 mL) will be frozen (−80°C) and stored by the Crimson Biobank until retrieval by the study investigators.

Two months post delivery, maternal and infant clinical data will be downloaded from the subject’s medical record and both discarded plasma samples will be retrieved from the Crimson Biobank for measurement of pGCD59. After retrieving the samples, we will break all links to the subject’s medical record and, thereafter, identify samples and data only through their code number without any access to the subjects’ personal information.

Patient and public involvement statement
Patients and the public were not involved in the design or the recruitment and conduct of this study, and will not be involved in the plans to disseminate the study results.

CLARITY REPORT
To efficiently download specified clinical information on both mother and infant from the EPIC medical record database, we have created a study-specific abstraction tool (Clarity report). Clarity is an EPIC native relational database warehouse made available to provide a platform for data analysis and volume data reporting. Most data from the EPIC real-time clinical database are downloaded to Clarity daily.

The Clarity system uses a variety of tools and algorithms to extract data attributes from the EMR. Examples of these methods and algorithms include:

- Extraction of problem lists and diagnoses using International Classification of Diseases categorisations.
- Extraction of problem lists using natural language processing.
- Extraction of data points from encounters across all Mass General Brigham health facilities (including both ambulatory and inpatient settings).
- Extraction of medications orders, imaging results and laboratory test results using dynamic date ranges based on prenatal/postpartum visits.
- Extraction of clinical data from different levels of the patient’s medical record including flowsheets, health maintenance encounters and procedure descriptions.
- Extraction of neonatal intensive care unit (NICU)-related activities by examining health level seven events.

A total of 213 variables will be obtained from the EMRs of the mothers and their infants. Table 1 presents a summary of the types of variables we will download from the EMR.

We have also created a Research Electronic Data Capture (REDCap) data import tool which we manually populate with the Clarity report. The data will be checked for accuracy, cleaned and imported into REDCap for storage or statistical analysis.

During the initial development and testing of the Clarity report we conducted data accuracy audits by reviewing pilot subjects’ EMRs. Similarly, we will conduct regular data audits for each set of extracted data prior to breaking the link to the participant’s personal identifiers.

The index test is levels of pGCD59 measured with the highly specific and sensitive ELISA assay described previously. The reference tests are the 50-g GLT and the 100-g 3-hour OGTT, respectively, administered as standard prenatal care to screen for and diagnose GDM.
Outcomes
The primary outcome is diagnosis of GDM at pregnancy weeks 24–28, extracted from the maternal medical record and adjudicated based on C&G criteria. A subject is classified as having GDM if she has at least two values at or above the following thresholds plasma glucose levels: fasting: ≥95 mg/dL, 1-hour 180 mg/dL, 2-hour 155 mg/dL and 3-hour 140 mg/dL. Subjects are also classified as having GDM if their 1-hour GTL result is ≥190 mg/dL.

Secondary outcomes include glucose intolerance (defined as one abnormal OGGT result) and adverse maternal and neonatal outcomes associated with GDM. Secondary maternal outcomes include a diagnosis of pre-eclampsia and caesarean section. Secondary neonatal outcomes include low birth weight as determined by the Ponderal Index (PI=fetal wt (g) × (100/fetal length (cm))^3), macrosomia (birth weight >4000 g), LGA (based on the Fenton curve), fetal malformations, including neural tube, congenital heart and abdominal wall defects, shoulder dystocia/birth injury, small for gestational age (SGA), neonatal hypoglycaemia, hyperbilirubinaemia and NICU admission.

Data analysis plan
The initial statistical analyses will involve generating a detailed description of the maternal and infant characteristics for the total cohort by GDM and GIGT status. Counts and percentages will be reported for categorical variables. For continuous variables, histograms will be generated to determine data distributions and identify potential outliers. Depending on data distribution we will report either means and SD or medians and IQRs.

Receiver operating characteristic curves will be used to evaluate the performance of pGCD59 in the first trimester to predict a diagnosis of GDM or GIGT. We will compute the sensitivity, specificity, percentage of correctly classified, and both positive and negative likelihood ratios for each of the percentile cut points of pGCD59. The area under the curve will be computed using the trapezoidal rule.

The associations between levels of pGCD59 and adverse maternal and infant outcomes will be modelled using linear or logistic regression, depending on the outcome, with adjustment for relevant confounders and covariates (figure 1). Linear regression will be used to determine the associations between pGCD59, birth weight and the PI. Additionally, logistic regression will be used to determine the associations between pGCD59 and pre-eclampsia, caesarean section, macrosomia, LGA/SGA, fetal malformations, intrauterine growth restriction, shoulder dystocia/birth injury, neonatal hypoglycaemia, hyperbilirubinaemia and NICU admission. Some potential confounders that will be considered are maternal age, race, BMI and gestational weight gain.

Missing data will be handled by assessing the pattern of missing and using multiple imputation if data are missing at random. The level for significance for all analyses will be set at p<0.05.

REFERENCES


